CLAIMS

1. A method to identify a preprocessing algorithm for raw data, the method comprising:

 providing an algorithm knowledge database including preprocessing algorithm data and feature set data associated with the preprocessing algorithm data;

 analyzing raw data to produce analyzed data;

 extracting from the analyzed data features that characterize the data;

 selecting a preprocessing algorithm using the algorithm knowledge database and features extracted from the analyzed data.

2. The method of claim 1 wherein the raw data comprises at least member selected from a group consisting of DSP data and IP data.

3. The method of claim 2 wherein:

 if the raw data comprises DSP data then the raw data is analyzed using at least one process selected from a group consisting or TFR-space transformation, phase map representation, and detection/clustering, and

 if the raw data comprises IP data then the raw data is analyzed using at least one process selected from a group consisting of detection/segmentation and region of interest shape characterization.

4. The method of claim 1 further comprising at least one member selected from a group consisting of

 data preparation and

 evaluating the selected preprocessing algorithm.

5. The method of claim 4 wherein the data preparation includes at least one member selected from a group consisting of conditioning/preprocessing, constant false alarm rate processing, and adaptive integration.

6. The method of claim 5 wherein the conditioning/preprocessing includes at least one member selected from a group consisting of interpolation, transformation, normalization, hardlimiting outliers, and softlimiting outliers.

7. The method of claim 4 further comprising the step of updating the algorithm knowledge base after evaluating the selected preprocessing algorithm.

8. A data mining system for identifying a preprocessing algorithm for raw data comprising:

at least one memory containing an algorithm knowledge database and raw data for processing;

random access memory having stored therein a computer program and which is coupled to the at least one memory such that the random access memory is adapted to receive:

at least one data analysis program to analyze raw data,

a feature extraction program to extract features from raw data, and

an algorithm selection program to identify a preprocessing algorithm.

9. The data mining system of claim 8 wherein the algorithm knowledge database and the raw data for processing are contained in a plurality of memories.

10. The data mining system of claim 8 wherein the data analysis program includes at least one member selected from a group consisting of a DSP data analysis program and an IP data analysis program.

11. The data mining system of claim 10 where

the DSP data analysis program is able to perform at least one subprogram selected from a group consisting of TFR-space transformation, phase map representation, and detection/clustering, and

the IP data analysis program is able to perform at least one subprogram selected from a group consisting of detection/segmentation and region of interest shape characterization.

12. The data mining system of claim 8 wherein the random access memory is also adapted to receive at least one member selected from a group consisting of a data preparation subprogram and an algorithm evaluation subprogram.

13. The data mining system of claim 12 wherein the data preparation program includes at least one member selected from a group consisting of a conditioning/preprocessing subprogram, a constant false alarm rate processing subprogram, and an adaptive integration subprogram.

14. The data mining system of claim 13 wherein the conditioning/preprocessing subprogram includes at least one member selected from a group that includes interpolation, transformation, normalization, hardlimiting outliers, and softlimiting outliers.

15. The data mining system of claim 12 wherein the algorithm evaluation program updates the algorithm knowledge database on the first storage device.

16. A data mining system for identify a preprocessing algorithm for raw data, the data mining system comprising

        a means for storing an algorithm knowledge database,

        a means for storing raw data;

        a means for data analysis on the raw data to produce analyzed data;

        a means for feature extraction from the analyzed data to produce a feature set;

        a means for algorithm selection using the feature set and the algorithm knowledge database.

17. The data mining system of claim 16 wherein the means for data analysis is selected from a group consisting of a means for DSP data analysis and a means for IP data analysis.

18. The data mining system of claim 17 wherein

        the means for DSP data analysis includes at least one member selected from a group consisting of a means for TFR-space transformation, a means for phase-map representation, and a means for detection/clustering, and

        the means for IP data analysis includes at least one member selected from a group consisting of a means for detection/segmentation and a means for region of interest shape characterization

19. The data mining system of claim 16 further comprising at least one member of a group consisting of:

a means for algorithm evaluation whereby the data mining system updates the algorithm knowledge database; and

a means for data preparation that converts the raw data into prepared data, wherein the means for data analysis operates on the raw data after it has been converted into the prepared data.

20. The data mining system of claim 19 wherein the means for data preparation includes at least one member selected from a group consisting of a means for conditioning/preprocessing of the raw data, a means for constant false alarm rate processing of the raw data, and a means for adaptive integration of the raw data.

21. The data mining system of claim 20 wherein the means for conditioning/preprocessing includes at least one member selected from a group consisting of a means for interpolation, a means for transformation, a means for normalization, a means for hardlimiting outliers, and a means for soft limiting outliers.

22. A data mining application comprising:

a) an algorithm knowledge database including preprocessing algorithm data and feature set data associated with the preprocessing algorithm data;

b) a data analysis module that is adapted to receive control of the data mining application when the data mining application begins;

c) a feature extraction module that is adapted to receive control of the data mining application from the data analysis module and that is available to identify a set of features; and

d) an algorithm selection module that is adapted to receive control from the feature extraction module and that is adapted to identify a preprocessing algorithm based upon the set of features identified by the feature extraction module using the algorithm knowledge database.

23. The data mining application of claim 22 wherein the algorithm selection module selects an algorithm from a group consisting of at least one DSP algorithm and at least one IP algorithm.

24. The data mining application of claim 23 wherein the algorithm selection module selects an algorithm using at least one member selected from a group consisting of energy compaction capabilities, discrimination capabilities, correlation capabilities.

25. The data mining application of claim 23 wherein

the algorithm selection module selects the at least one DSP algorithm if and only if the data analysis module uses at least one member of a group consisting of a short-time Fourier transform coupled with linear predictive coding analysis, a compressed phase-map representation, and a detection/clustering process; or

the algorithm selection module selects the at least one IP algorithm if and only if the data analysis module uses at least one member of a group consisting a procedure operable to provide at least one a region of interest by segmentation, a procedure to extract local shape related features from a region of interest; a procedure to extract two-dimensional wavelet features characterizing a region of interest; and a procedure to extract global features characterizing all regions of interest

26. The data mining application of claim 25 wherein the detection/clustering process includes at least one member selected from a group consisting of (a) an expectation maximization algorithm and (b) procedures that perform operations of setting a hit detection threshold, identifying phase-space map tiles, counting hits in each identified phase-space map tile, and detecting the phase-space map tiles for which the hits counted exceeds the hit detection threshold.

27. The data mining application of claim 22 further comprising at least one member of a group consisting of:

an advanced feature extraction module available to receive control from the algorithm selection module and to identify more features for inclusion in the set of features;

a data preparation module that is available to receive control after the data mining application begins, wherein the data analysis module is available to receive control from the data preparation module; and

an algorithm evaluation module that evaluates performance of the preprocessing algorithm identified by the algorithm selection module and updates the algorithm knowledge database.

28. The data mining application of claim 27 wherein the data preparation module includes at least one member selected from a group consisting of a conditioning/preprocessing process, a constant false alarm rate processing process to identify and extract long term trend lines, and an adaptive integration process.

29. The data mining application of claim 28 wherein

the conditioning/preprocessing process includes at last one member selected from a group consisting of interpolation, transformation, normalization, hardlimiting outliers, and softlimiting outliers; and

the adaptive integration includes at least one member selected from a group consisting of subspace filtering and kernel smoothing.

30. A data mining product embedded in a computer readable medium, comprising:

at least one computer readable medium having an algorithm knowledge database embedded therein and having a computer readable program code embedded therein to identify a preprocessing algorithm for raw data, the computer readable program code in the computer program product comprising:

computer readable program code for data analysis to produce analyzed data from the raw data;

computer readable program code for feature extraction to identify a feature set from the analyzed data; and

computer readable program code for algorithm selection to identify a preprocessing algorithm using the analyzed data and the algorithm knowledge database.

31. The data mining product of claim 30 wherein the data mining product is embedded in a plurality of computer readable media.

32. The data mining product of claim 30 wherein the computer readable program code for data analysis includes at least one member selected from a group consisting of computer readable program code for DSP data analysis and computer readable program code for IP data analysis.

33. The data mining product of claim 32 wherein

the computer readable program code for DSP data analysis includes at least one member of a group consisting of computer readable program code for TFR-space transformation, computer readable program code for phase map representation and computer readable program code for detection/clustering, and

the computer readable program code for IP data analysis includes at least one member of a group consisting of computer readable program code for detection/segmentation, and computer readable program code for region of interest shape characterization.

34. The data mining product of claim 30 further comprising at least one member selected from the group consisting of

computer readable program code for data preparation to produce prepared data from the raw data, wherein the computer readable program code for data analysis operates on the raw data after it has been transformed into the prepared data; and

computer readable program code for algorithm evaluation to evaluate the preprocessing algorithm selected by the computer readable program code for algorithm selection.

35. The data mining product of claim 34 wherein the computer readable program code for algorithm evaluation is operable to modify the algorithm knowledge database.

36. The data mining product of claim 34 wherein the computer readable program code for data preparation includes at least one member from a group consisting of computer readable program code for conditioning/preprocessing, computer readable program code for constant false alarm rate processing, and computer readable program code for adaptive integration.

37. The computer program product of claim 36 wherein the computer readable program code for conditioning/preprocessing includes at least one member selected from a group consisting of computer readable program code for interpolation, computer readable program code for transformation, computer readable program code for normalization, computer readable program code for hardlimiting outliers, and computer readable program code for softlimiting outliers.